DOCUMENT RESUME

ED 354 250 TM 019 478

AUTHOR Kim, JinGyu
TITLE Three Approaches for the Integration of Teaching, Testing, and Learning.
PUB DATE Nov 92
NOTE 21p.; Paper presented at the Annual Meeting of the Mid-South Educational Research Association (21st, Knoxville, TN, November 11-13, 1992).
PUB TYPE Information Analyses (070) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Adaptive Testing; Behaviorism; Cognitive Psychology; *Computer Assisted Testing; Educational Theories; Elementary Secondary Education; Feedback; Higher Education; Holistic Approach; *Integrated Activities; Knowledge Level; *Learning; *Teaching Methods; *Test Use
IDENTIFIERS Teaching to the Test

ABSTRACT

The relationships among teaching, testing, and learning are examined and reconceptualized based on different theories and computer adaptive testing. Approaches that integrate teaching, testing, and learning include: (1) behavioral psychology; (2) cognitive psychology; and (3) computer adaptive testing (CAT). Some behaviorists have assumed that all important learning objectives can be specified and measured completely, and some behaviorists have argued that teaching to tested objectives is synonymous with good instruction. Some cognitive psychologists have argued that the mechanisms for children's acquisition of knowledge are linked intimately with cognitive theory, and that assessment of knowledge acquisition must be integrated with the instructional process. The diagnostic approach and prior knowledge approach are considered as the main techniques of integrating teaching and testing. Proponents of CAT argue that it can integrate teaching and testing as it provides rapid feedback about performance. All three approaches identify the value and importance of linking instruction and testing. There is a 67-item list of references. (SLD)

THREE APPROACHES FOR THE INTEGRATION OF TEACHING, TESTING, AND LEARNING

JinGyu Kim

The University of Alabama

A Paper Presented at the Annual Meeting

of the

Mid-South Educational Research Association

Knoxville, Tennessee

November 11-13, 1992

I. INTRODUCTION

The use of tests is increasing because they serve important functions in education. Educational tests are used in schools to achieve four purposes: accountability, selection, evaluation of educational innovations and projects, and instructional guidance (Tyler & White, 1979). Above all, Tyler and White indicate that instructional guidance is the educational activity which is least served by existing tests. The appropriate use of tests for instructional guidance is open to debate, but an easily overlooked point is that this purpose has more to do with the bureaucratic nature of schools than real teaching and individual students' learning. Typically, students receive only an aggregate score after several weeks delay. Individual students cannot get rapid and specific feedback on how well they have done. Moreover, most standardized tests have weak relationships to the teaching-learning process. Such tests cannot provide optimal information to guide teachers in their choice of topics and skills.

A systematic investigation of the relationships among the major components of teaching, testing, and learning is needed. Much attention has been paid to teaching, testing, and learning as individual constructs. The substantial interrelationships of the three major components of schooling have relatively been ignored. How do students' conceptions of learning influence their conceptions of teaching and testing? How do teaching behaviors affect student's conceptions of learning and their perceptions about testing? How do testing procedures relate to student's conceptions of learning and their expectancies about teaching? How do a variety of learning theories such as behaviorism, connectionism, Gestalt psychology, constructivism, and brain functional theory enable us to see the relationships among teaching, testing, and learning in a new way? There are many ways of thinking about how the three major components relate to each other. For example, two way

connections such as the teaching-learning relationships and three way connections such as the teaching-testing-learning interrelationships can be examined. In this paper the following three approaches for integrating teaching, testing, and learning will be explored: behavioral psychology, cognitive psychology, and computer adaptive testing (CAT). The purpose of this paper is (a) to examine the relationships among teaching, testing, and learning and (b) to reconceptualize these relationships based on different theories and computer adaptive testing.

II. TESTING, TEACHING, AND LEARNING

As core constructs of schooling, teaching, testing, and learning have formed their own unique, various, and systematic theories individually. The underlying postulate under consideration in this paper is that there are some ways of thinking about how the three major constructs relate to each other.

The teaching-learning process has been explored depending on how learning theory relates to teaching practices. Some learning theorists (e.g. Skinner, 1954, 1958) believe that the learning principles drawn from laboratory research can be applied directly to teaching practices, while others (e.g. Bruner, 1966; Snelbecker, 1974) have raised the need for an instructional theory that presents more specific and prescriptive principles to classroom teachers.

With regard to the testing-learning connection, there has been much theoretical discussion of the value of testing in the learning process but only spasmodic research in the area (Balch, 1964; Kirkland, 1971; Gaynor & Millham, 1976; Rudman et al., 1980; Halpin & Halpin, 1982; Nance, 1991). Crooks (1988) summarizes the results of 14 specific fields of research that cast light on the relationships between classroom evaluation practices and student learning outcomes. He concludes that classroom evaluation has powerful direct and indirect impacts, which may be positive or negative, and thus deserves very thoughtful planning and implementation.

Recently, the linkage between testing and teaching or instruction has been raised as a special issue on how students promote their learning (Tyler & White, 1979; Airasian & Madaus, 1983; Burstein, 1983; Linn, 1983; Fielding & Schalock, 1985; Fielding & Shaughnessy, 1986; Glaser, 1986; Fielding, 1987). Although learning can occur without teaching or testing, each of these have been accepted as one of the important variables in the learning process.

On the other hand, the three way connections integrating teaching, testing, and learning can be conceptualized by a variety of criteria: teacher control (other-regulation) versus learner control (self-regulation), sequential learning tasks versus knowledge structure, and learning outcome versus learning process. Based on the above criteria, the three way connections can be divided into behavioral psychology approach, cognitive psychology approach, and computer adaptive testing (CAT) approach.

The behavioral psychology approach promotes teacher control, sequential learning hierarchies, other-regulation, and learning outcomes. In Skinner's view, learning is a change in behavior represented by increases in response frequency (Gredler, 1992). Learning is seen to be linear and sequential. Complex learning can occur only by the accretion of elemental, prerequisite learning (Gagne, 1968). Testing should be closely allied with teaching to facilitate learning. Thus, the test is corrected immediately, and if students achieve the predetermined learning objectives they move on to a new learning task assigned by the teacher. This approach emphasizes the link between teaching and testing in a three way connection. Thus, testing activities are considered as part of instruction. For example, programmed instruction (Skinner, 1958), mastery learning (Bloom, 1971), computer assisted instruction (CAI), and individually prescribed instruction (Education U.S.A., 1968) apply testing techniques as an essential component in the program. In this approach teaching aspects related to testing are considered in the learning process.

The cognitive psychology approach focuses on learner control, knowledge structure, active self-regulation, and learning process. In this approach learning is not linear and is not acquired by assembling bits of simple facts. Rather, learning is defined as constructive internalization of cognitive maps (Tolman, 1932), schemas (Rumelhart & Ortony, 1977), and other kinds of external knowledge functionally organized cognitive process (Luria, 1980), as constructive, dynamic, and creative reconceptualization of internal knowledge (Iran-Nejad, 1990), or as connection weights modification among network elements (Bereiter, 1991). The major principle of the cognitive approach is that learning occurs under the active control of one internal source of self-regulation (Atkinson & Shiffrin, 1968) or under the active and dynamic sources of internal self-regulation (Iran-Nejad, 1990). The cognitive psychology approach highlights more latent and internal aspects of learning in the three way connection than do the behavioral psychology approach. Therefore, this approach pays much more attention to learning aspects related to teaching and testing in the learning process.

The computerized adaptive testing (CAT) approach shares elements with both behavioral psychology and cognitive psychology approaches. Learner control and active self-regulation might work better than teacher control and other-regulation in the CAT. Thus, the computer and teacher would be assistants to the student's self learning. The teacher's role is that of a coach rather than instructor. In a CAT, tests items can be interpreted as the squential learning tasks in that the items would be presented squentially, based on the students' knowledge state or ability. The CAT approach emphasizes both learning process and learning outcome in that its goal is to optimize the learning process and the test item serves as a perfect example of the learning objectives. Thus, this approach is called an "ecletic" approach in a three way connection. Unlike two traditional approaches, the CAT approach can easily integrate the three core constructs of schooling. Because the computer can play an important role in integrating teaching, testing, and learning. Through the

use of computer technology we can design an integrated testing and learning environment. The concept of teaching in a CAT refer to the design of as integrated testing-learning environment. The behavioral psychology approach focuses on teaching aspects and do the cognitive psychology approach on learning aspects, while this approach highlights testing aspects related to student's self-learning environment with the computer and teacher.

III. APPROACHES INTEGRATING TEACHING, TESTING, AND LEARNING

Behavioral Psychology Approach

Behaviorism has formed an important theoretical core between the 1920s and the 1960s. At that time the classical psychometric theory was well suited for compatibility with behavioral theories of learning. This had implications for test construction and led to the kinds of standardized group achievement and intelligence tests in predominant use in schools today (Wiilson, 1991).

There are many learning theories of behavioral psychology which emphasize the testing and instructional principles (Skinner, 1954; Bloom, 1956; Education U.S.A., 1968; Gagne, 1970; Bloom,1971; Resnick, Wang, & Kaplan, 1973; Popham, 1978). Whether coached in terms of teaching machines, learning hierarchies, programmed learning, mastery learning, or criterion-referenced testing, these authors share the same learning theory, behavioral psychology.

Shepard (1991) presented an insightful analysis of the origins of measurement specialists' learning theory in programmed instruction and behavioral psychology. He discussed the criterion-referenced testing paradigm which is grounded in the behavioral learning theories. In this paradigm, tests should exactly specify desired behavioral outcomes of instruction and should be used at each learning step. We can say that the role of testing is to assess the mastery level of the

various objectives, to diagnose progress in student learning unit tasks, and to provide feedback to teachers and students on what has been and needs to be learned.

The behavioral learning theory has relied on the true score theory. Theoretically, a person's score is conceived as having two components, true score and error score. From several simple assumptions the item characteristics such as reliability and validity are produced. The total score and the right or wrong scoring have played an important role in item analysis. Test items also are constructed using a model in which items are selected from templates of objectives or content lists and then are either varied in difficulty or held at constant difficulty to shape overall psychometric characteristics of the test (Willson, 1989).

It should be clear that the behaviorists and programmed learning model also make assumptions about the nature of tests. In understanding the nature of a test, they encourage the use of external predictors based on total scores and the right or wrong answers, while they pay little attention to the semantic, structural, and visual features of the items themselves which can estimate the examinees' ability or achievement. Each of the learning steps is small enough that highly homogeneous tests can be used to measure mastery at each step without inference to some broader set of test questions or criterion performances (Shepard, 1991). Therefore, it should be clear from the behaviorist perspective that observable learning tasks and learning objectives are equivalent and that teaching to tested objectives is synonymous with good instruction which leads to students' learning effectively.

However, the most serious problems from behavioral perspectives is that higher order skills are not introduced until after prerequisite skills have been mastered. Resnick and Resnick (1991) explained the inadequacies of associationist and behaviorists theories. They described the assumptions of decomposability and decontextualization. They assume that component skills can

be adequately defined and mastered independently and out of context. Only then are more advanced thinking skills acquired by "adding up" or assembling component abilities (Shepard, 1991).

Cognitive Psychology Approach

Cognitive psychology is concerned with various mental abilities such as perception, learning, memory, reasoning, problem solving, and decision making and it represents the mainstream of thinking in both psychology and education. Neisser (1967) defined cognitive psychology as the study of the cognitive structures mechanisms, and elaborated that it has a right to expect some discussion of thinking, concept-formation, remembering, problem-solving, and the like.

Many psychologists interested in the interface between cognitive psychology and psychometric testing believe that cognitive theories can be related to testing practices (Curtis & Glaser, 1983; Sternberg, 1984, 1991; Glaser, 1984, 1986; Willson, 1989). There are some techniques combining testing and teaching from the cognitive perspectives. Tyler and White (1979) suggest a diagnostic model, a psychological model, background knowledge model, and a psycho-linguistic model. Glaser (1986) presents the analysis of rule of performance, assessment of prior knowledge, the coordination of basic and advanced performance, and the nature of competence and expertise. Willson (1989) recommends four primary components of ... (1) prior knowledge required for solution, (2) intened and unintended learner processes, (3) item processing demands, and (4) response qualities to be assessed (response characteristics and latency). In this paper the diagnostic approach and prior knowledge technique will be discussed below.

The diagnostic approach is considered as the main technique of integrating testing and teaching in the above mentioned research (diagnostic model, Tyler & White, 1979; the analysis of rule of performance, Glaser, 1986; and response quality, Willson, 1989). This approach has

focused on error analysis to explain, represent, and measure students' misconceptions in mathematics, science, and language arts. First, Brown and Nurton (1978) developed a computer program "BUGGY" that can diagnose students' erroneous algorithms (rule of operation) resulting from misconceptions (bugs) in whole number subtraction problems. Their study was given an impetus that attempted to link incorrect responses to arithmetic items to instructional deficiency and consequent instructional strategies. Tatsuoka (1983) and her colleagues (Tatsuoka & Linn, 1983; Tatsuoka & Tatsuoka, 1987; Tatsuoka, Birenbaum, & Arnold, 1989) also developed a computer program that can diagnose a number of erroneous rules in sign-number addition and substraction problems. The selection of options in multiple choice items becomes much more important when response qualities to answers on test items is considered. The construction of the options now requires formal analysis of knowledge and processing demands made by each option on the test taker. The cognitive research on error analysis can provide useful information in evaluating instruction or instructional materials as well as specific prescriptions for planning remediation for a student. The cognitive research on error analysis provides a new door in understanding the teaching-testing-learning connection. In a cognitive approach, tests may be informative even if the classical reliability is low or even zero; score pattern consistency is valued as highly as between-subject variance consistency. Error in this framework is linked to knowledge state and to instructional procedures that are intended to improve the students' domain and strategic knowledge (Willson, 1991). There is recent evidence (Birenbaum & Tatsuoka, 1987) supporting what many have long suspected, that different processes are required and used for the multiple choice and open ended types of items. Both cognitive and developmental demands of various item formats should be reconsidered and examined from an information processing perspective (Willson, 1989).

The prior knowledge approach has stemmed from the traditional concept of transfer. The newer cognitive concern for the role of prior knowledge in learning recognizes that for meaningful forms of learning this process is more complex than the one suggested by earlier approaches to transfer (Shuell, 1986). This prior knowledge has played an important role in the acquisition of new knowledge. Not surprisingly, students who know more about a topic or domain, understand and remember content better than those who have a limited background in the domain (Chi, 1985). It is well known that students attempt to understand and think about new information in terms of what they already know (Bransford & Franks, 1976; Siegler & Klahr, 1982; Siegler, 1983). Moreover, high levels of learning and understanding can be fostered by insuring contact between new information and the student's prior knowledge, which then can be restructured through instruction. We know that students often begin learning with some common misconceptions about the scientific phenomena (e.g. Champagne, Kloper, & Gunstone, 1982) and that remnants of these misconceptions even persist in students who receive high grades in a course (Gunstone & White, 1981). These misconceptions represent what students consider to be appropriate ways of dealing with the problem, given their current knowledge structure, namely prior knowledge. Thus, the assessment of prior knowledge provides teachers with useful insights into the type of instruction and an awareness of the cognitive processes that must be used by the learner. Willson (1989) indicates that the intended prior knowledge necessary to answer an achievement item might be organized by the expected connections or links between information provided in the stem, prior domain knowledge resident in the learner, and the correct answer. As Birenbaum and Tatsuoka (1987) pointed out, the form of the question, open-ended or multiple choice, may alter the connections selected by the respondent as well as alter the control process and information processes used. In psychometric theory, prior knowledge is assumed to be the mean of the difficulty level of the test (Willson, 1991). There is an appeal to developmental research and educational acumen to locate

the test at an appropriate difficulty level, and it is then fine tuned by an adjustment to the items included in a test.

The cognitive research on prior knowledge opens to a new perspective in understanding the teaching-testing-learning connection. First, teachers must know how to get students actively engaged in learning activities. In other words, teachers need to know more about the way in which specific content and instructional procedures engage the cognitive structure and knowledge structure.

Second, prior knowledge can provide students with the internal sources of self learning. They can get the information that needs to be known and information about the conditions under which use of this knowledge is appropriate (Glaser, 1986).

Third, test developers can get the information for maintaining an appropriate difficulty level in test construction when they can measure the test taker's prior knowledge.

The cognitive psychology approach emphasizes the learner's control, active and dynamic sources of internal self-regulation, knowledge srtructure, and learning processes. Above all, the cognitive psychology approach left a new door open in understanding the teaching-testing-learning connection.

Computer Adaptive Testing Approach

An adaptive testing is an alternative procedure that matches different sets of test items to different examinees' previous responses to items or abilities during the administration of a test. In an adaptive test, correct responses are followed by the presentation of more difficult items, and incorrect responses are followed by the presentation of easier items. Performance on initial items provides evidence of examinee ability that indicates what the subsequent item should be. The test is terminated when a previously determined criterion is reached. Adaptive testing has also been

referred to as tailored, response-contingent, programmed, individualized, branched, and sequential testing (Weiss, 1985).

The earliest application of adaptive testing was in the work of Alfred Binet on intelligence testing in 1908 (Weiss, 1985). Later, examples of adaptive tests were two-stage testing (Angoff & Huddleston, 1958), pyramidal testing (Krathwohl & Huyser, 1956), and the flexilevel test (Lord, 1971). Recently, adaptive tests are administered by an interactive computer, hence computerized adaptive testing (CAT). A computer algorithm is used to match the difficulty levels of the items administered to the ability level of each examinee. Most CATs are based on item response theory (IRT) methodologies (Wainer, 1983; Weiss, 1982, 1985). IRT measures the empirical difficulty of each item, estimates the accuracy of the ability, and decides whether the test is over. Thus, the procedures of adaptive testing are somewhat different from the conventional testing in which all examinees are administered the same fixed set of items. The common procedures of CATs (Weiss & Kingsbury, 1984; Reckase, 1989; Hambleton, Zaal, & Pieters, 1991) are summarized as follows; (1) building the item bank, (2) selecting the items, (3) scoring and ability estimation, and (4) stopping the test.

The CATs have resulted not only in new forms of testing, but in resources that help schools and teachers store, manage, and use test information. Several software packages in CAT are now available for scoring criterion-referenced tests and reporting test results for specific learning objectives. When the potential benefits of computers in testing become realized, CATs will become integrated computer-based systems of instruction, testing, and learning. CAT makes it feasible to monitor individual progress more frequently, accurately, and comprehensively. With a thorough, continuing assessment of what each student needs to learn, CAT could be a typical model of the integration of teaching, testing, and learning.

Several learning principles can be drawn from CAT to integrate the teaching-testing-learning connection. First, the immediate feedback principles can be found in CAT. The responses of each item are followed by the rapid feedback. This principle is matched to that of programmed instruction which prompts the learner to emit a target behavior and then reinforce that behavior by confirming it (Skinner,1954, 1968). This principle has survived the transition from a behavioristic to cognitive learning paradigm (Shuell, 1986). The cognitive view of learning equally emphasizes the consequences of behavior, but it sees their operation as more than the mere strengthening of automatic stimulus-response associations. In this view, feedback serves as sources of new information necessary for verification of retrieval accuracy, concept development, skill refinement, and metacognitive adaptation (Bangert-Drowns, Kulik, Kulik, & Morgan, 1991).

Second, prior knowledge has played an important role in CAT. Item selections are based on the performance on initial items. Two different explanations related to learning principles are inferred from responses to items. Suppose that each set of test items serves as a perfect instantiation of the learning objectives. The previous responses to initial items have the same feature of prerequisite learning from the behavioral perspectives. If prior knowledge is acquired by the initial items and new information is given by the subsequent items, the prior knowledge contributes to student learning as the internal source that is evident in research on cognitive learning.

Third, CAT requires the active response from the learner. In a CAT, students perceive the test items from CRT screen and respond actively to the stem and options of a given item. The active learning principle is based on the cognitive approach that learning is an active, constructive, and goal-oriented process which is dependent upon the mental activities of the learner (Shuell, 1986). This principle is also matched to Vygotsky's active adaptation which humans actively adapt to the environment (Gredler, 1992).

Fourth, substantial students' learning in CAT can be related to the internal self-regulation processes of the learner (Iran-Nejad, 1992). In a CAT, students choose the correct answers by external self-regulation such as information provided in the stem, active self-regulation such as prior knowledge resident in the learner, and dynamic self-regulation such as postdiction on their expected correct answers.

Finally, the CAT can open a new door for the optimization of learning. The goal of the CAT is to optimize the testing-learning process. This will require the continous measurement of the person by item interaction. This principle is similar to that of aptitude treatment interaction (ATI, Cronbach & Snow, 1977).

IV. SUMMARY AND CONCLUSION

This paper has presented discussions of three approaches that have played an important role in the integration of teaching, testing, and learning. Some behaviorists have assumed that all important learning objectives could be specified and measured completely. Each of the learning steps is small enough that highly homogeneous tests can be used to measure mastery at each step. Furthermore, behaviorists have argued that teaching to tested objectives is synonymous with good instruction.

Some cognitive psychologists have argued that the mechanisms for children's acquisition of knowledge are intimately linked with cognitive theory, and assessment of knowledge acquisition must be integrated with the instructional process. The diagnostic approach and the prior knowledge approach are considered as the main techniques of integrating testing and teaching.

The proponents of the computer adaptive testing approach have argued that CAT cannot only integrate both instructional and testing sequence in the same computer-delivered tutorial

lesson, but provide rapid feedback about performance to teachers and students based on the students' previous performance.

In conclusion, the value and importance of linking instruction and testing in the learning process are identified by the three prominent approaches. From behavioristic perspective, the use of learning objectives and equivalent test items can be not only monitored students' progress, but also led to good instruction. Tests based on cognitive perspectives illuminate previous hidden aspects of student thinking and performance such as prior knowledge model and diagnostic models. Finally, as the potential benefits of the computer in testing become realized, the CAT over the conventional tests will soon become a popular and typical model in that it can provide students with an integrated self-regulated learning and testing environment.

REFERENCES

Airasian, P.W., & Madaus, G.F. (1983). Linking testing and instruction: Policy issues. Journal of Educational Measurement, 20(2), 103-118.

Angoff, W.H., & Huddleston, E.M. (1958). The multi-level experiment: A study of a two-level test system for the College Board Scholastic Aptitude Test. (Statistical Report No. SR-58-21). Princeton, NJ: Educational Testing Service.

Balch, J. (1964). The influence of the evaluating instrument on students' learning. American Educational Administration and Supervision, 1, 169-182.

Bangert-Drowns, R.L., Kulik, C., Kulik, J. A., & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. Review of Educational Research, 61(2), 213-238.

Bereiter, C. (1991). Implications of connectionism for thinking about rules. Educational Researcher, 20(3), 10-16.

Bloom, B.S.(ed.). (1956). Taxanomy of educational objectives: Handbook I, cognitive domain. New York: David McKay.

_____(1971). Mastery learning. In J.H. Block (ed.), Mastery learning: Theory and practice(pp. 47-63). New York: Holt, Rinehart, & Winston.

Bruner, J.S.(1966). Toward a theory of instruction. Cambridge, MA: Harvard University Press.

Brown, J.S., & Burton, R.R. (1978). Diagnostic models for procedual bugs in basic mathematical skills. Cognitive Science, 2, 155-192.

Burstein, L.(1983). A word about this issue. Journal of Educational Measurement, 20(3 & 4), 99-101.

Champagne, A.B., Klopper, L.E., & Gunstone, R.F. (1982). Cognitive research and the design of science instruction. Educational Psychologist, 17, 31-53.

Chi, M.T.H.(1985). Interactive roles of knowledge and strategies in the development of organized sorting and recall. In S.F. Chipman, J.W. Segal, & R. Glaser (Eds.), Thinking and learning skills(vol.2, pp. 457-484). Hillsdale, NJ: Erlbaum.

Crooks, T.J. (1988). The impact of classroom evaluation practices on students. Review of Educational Research, 58(4), 438-481.

Curtis, M.E., & Glaser, R. (1983). Reading theory and the assessment of reading achievement. Journal of Educational Measurement, 20(2), 133-147.

Education U.S.A.(1968). Individually prescribed instruction. Washington, DC: Author.

Fielding, G.D. (1987). Trends in integrating teaching and testing. ERIC Digest. Princeton, NJ: ERIC clearinghouse on tests, measurement, and evaluation. (ERIC No. ED 284 912).

Felding, G.D., & Schalock, H.D. (1985). Integrating teaching and testing: A handbook for high school teachers. Monmouth, OR: Oregon State System of Higher Education, Teaching Research Division. (ERIC No. ED 257 821).

Felding, G.D., Shaughnessy, J., & Duckworth, K. (1986). Effects on teacher practice of a staff development program for integrating teaching and testing. Eugene, OR: Center for Educational Policy and Management, University of Oregon. (ERIC No. ED 267 501).

Gagne, R.M. (1970). The conditions of learning (2nd ed.). New York: Holt, Rinehart, & Winston.

_____ (1968). Learning hierarchies. Educational Psychologist, 6, 1-9.

Gaynor, J., & Millham, J. (1976). Student performance and evaluation under variant teaching and testing methods in a large college course. Journal of Educational Psychology, 68, 312-317.

Glaser, R. (1986). The integration of instruction and testing. The redesign of testing for the 21st century: Proceedings of the 1985 ETS invitational Conference. Princeton, NJ: Educational Testing Service.

_____ (1984). Education and thinking: The role of knowledge. American Psychologist, 39, 93-104.

_____ (1981). The future of testing: A research agenda for cognitive psychology and psychometrics. American Psychologist, 36, 923-936.

Gredler, M.E. (1992). Learning and instruction: Theory into practice (2nd ed.). New York: Macmillan Publishing Company.

Gunstone, R.F., & White, R.T. (1981). Understanding of gravity. Science Education, 65, 291-300.

Halpin, G., & Halpin, G. (1982). Experimental investigation of the effectss of study and testing on student learning, retention, and ratings of instruction. Journal of Educational Psychology, 74(1), 32-38.

Iran-Nejad, A. (1990). Active and dynamic self-regulation of learning processes. Review of Educational Research, 60, 573-602.

Iran-Nejad, A., & Cecil, C. (1992). Interest and learning: A biofuntional perspective. In K.A. Renninger, S. Hidi, & A. Krapp (Eds.), The role of interest in learning and motivation. Hilsdale, NJ: Erlbaum.

Kirkland, M.C. (1971). The effects of tests on students and schools. Review of Educational Research, 41(3), 303-350.

Krathwohl, D.R., & Huyser, R.J. (1956). The sequential item test (SIT). American Psychologist, 2, 419.

Linn, R. (1983). Testing and instruction: Links and distinctions. Journal of Educational Measurement, 20(2), 179-189.

Lord, F.M. (1971). The self-scoring flexilevel test. Journal of Educational Measurement, 8, 147-151.

Luria, A.R. (1980). Higher cortical functions in man (2nd ed.). New York: Wiley.

Michael, L.J. (1988). Simple and effective algorithms: Computer-adaptive testing. Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 5-9, 1988). (ERIC No. ED 294 918).

Nance, J.L. (1991). Testing and learning. Journal of Instructional Psychology, 18(3), 205-207.

Neisser, U. (1967). Cognitive psychology. New York: Appleton-Century-Crofts.

Olsen, J.B. (1990). The four generations of computerized testing: Toward increased use of AI and Expert Systems. Educational Technology, 30(3), 36-41.

Popham, W.J. (1978). Criterion-referenced measurement. Englwood Cliffs, NJ: Prentice-Hall.

Reckase, M.D. (1989). Adaptive testing: The evolution of a good idea. Educational Measurement: Issues and Practice, 8(3), 11-15.

Resnick, L.B.,& Resnick, D.P. (1991). Assessing the thinking curriculum: New tools for educational reform. In B.R. Gifford & M.C. O'Connor (Eds.), Changing assessments: Alternative views of aptitude, achievement, and instruction. Boston: Kluwer Academic Publishers.

Resnick, L.B., Wang, M.C., & Kaplan, J. (1973). Task analysis in curriculum design: A hierarchically sequenced introductory mathematics curriculum. Journal of Applied Behavioral Analysis, 6, 679-710.

Rudman, H.E., Kelly, J.L., Wanous, D.S., Mehrens, W.A., Clark, C.M., & Porter, A.C. (1980). Integrating assessment with instruction: A review (1922-1980) (Research Series No. 75). East Lansing, MI: Michigan State University, Institute for Research on Teaching. (ERIC Document Reproduction Service No. ED 189 136).

Rumelhart, D.E., & Ortony, A. (1977). The representation of knowledge in memory. In R. C. Anderson, R.J. Spiro, & W.E. Montague (Eds.), Schooling and the acquisition of knowledge (pp. 99-136). Hillsdale, New Jersey: Erlbaum.

Shepard, L.A. (1991). Psychometricians' beliefs about learning. Educational Researcher, 20(6), 2-16.

Shuell, T.J. (1986). Cognitive conceptions of learning. Review of Educational Research, 56(4), 411-436.

Siegler, R.S. (1983). Five generations about cognitive development. American Psychologist, 38, 263-277.

Siegler, R.S., & Klahr, D. (1982). When do children learn? The relationship between existing knowledge and the acquisition of new knowledge. In R. Glaser (Ed.), Advances in instructional psychology(vol. 2, pp. 121-211). Hillsdale, NJ: Erlbaum.

Skinner, B.F. (1954). The science of learning and the art of teaching. Harvard Educational Review, 24, 86-97.

_____ (1958). Teaching machines. Science, 128, 967-97.

_____ (1968). The psychology of teaching. Englewood Cliffs, New Jersey: Prentice-Hall.

Snelbecker, G.E. (1974). Learning theory, instructional theory, and psychoeducational design. New York: McGraw-Hill.

Sternberg, R.J. (1984). What cognitive psychology can (and cannot) do for test development. In B.S., Plake (ed.). Social and technical issues in testing: Implications for test construction and usage. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

_____(1991). Cognitive theory and psychometrics. In R.K., Hambleton & J.N. Zaal (eds.), Advances in educational and psychological testing: Theory and applications. Boston, Mass.: Kluwer Academic Publishers.

Tatsuoka, K.K. (1983). Rule space: An approach for dealing with misconceptions based on item response theory. Journal of Educational Measurement, 20(4), 34-38.

Tatsuoka, K.K., & Linn, R.L. (1983). Indices for detecting unusual response patterns: Links between two general approaches and potentials applications. Applied Psychological Measurement, 7( ), 81-96.

Tatsuoka, K.K., & Tatsuoka, M.M. (1987). Bug distribution and statistical pattern classification. Psychometrika, 52( ), 193-206.

Tatsuoka, K.K., Birenbaum, M., & Arnold, J. (1989). On the stability of students' rules of operation for solving arithmetic problems. Journal of Educational Measurement, 26(4), 351-361.

Tolman, E.C. (1932). Purposive behavior in animals and men. New York: Appleton-Century-Crofts.

Tyler, R.W., & White, S.H. (1979). Testing, teaching and learning: Report of a conference on research on testing August 17-26, 1978. Washington, D.C.: U.S. Department of Health, Education, and Welfare and the National Institute of Education.

Wainer, H. (1983). On item response theory and computerized adaptive tests. The Journal of College Admissions, 27(4),9-16.

Weiss, D.J. (1985). Adaptive testing by computer. Journal of Consulting and Clinical Psychology, 53(6), 774-789.

_____ (1982). Improving measurement quality and efficiency with adaptive testing. Applied Psychological Measurement, 6, 473-492.

Weiss, D.J., & Kingsbury, G.G. (1984). Bias and information of Bayesian adaptive testing. Applied Psychological Measurement, 8, 273-285.

Willson, V.L. (1989). Cognitive and developmental effects on item performance in intelligence and achievement tests for young children. Journal of Educational Measurement, 26(2), 103-119.

_____ (1991). Performance assessment, psychometric theory and cognitive learning theory: Ships crossing in the night. Contemporary Education, 62(4), 250-254.

Wise, S.L., & Plake, B.S. (1989). Research on the effects of administering tests via computers. Educational Measurement: Issues and Practice, 8(3), 5-10.